

Developing a Trilingual Dictionary for Islamic Boarding School Students: A Corpus-Based Method

The development of trilingual dictionaries utilizing corpus data has gained traction as a pivotal method in enhancing language learning. However, the creation of a trilingual dictionary for English, Arabic, and Indonesian, particularly in the context of Islamic boarding schools, has been infrequent. Such development is deemed necessary to address regulations that prohibit students from searching for words online. To systematically create a trilingual dictionary that authentically reflects language use, this study employed a 4-D design methodology: defining, designing, developing, and disseminating. This process involved identifying the needs of 50 students and five English teachers from Darul Muttaqien NWDI Perian through questionnaire, interview, and focus group discussion (FGD). The findings revealed that students face significant lexical gaps in both general and pesantren-specific English terminology, primarily due to limited language exposure and practice opportunities. Therefore, the development of a trilingual dictionary was essential following the results of the needs analysis. This study extended beyond the immediate benefits of vocabulary enhancement but also suggested a paradigm shift in how trilingual dictionaries can be constructed and utilized in educational contexts, especially in Islamic boarding schools. By leveraging corpus data, teachers can create more dynamic and responsive language resources that cater to the evolving needs of students.

Keywords: Islamic trilingual dictionary, vocabulary enhancement, corpus-based methodology, 4-D design

Entwicklung eines dreisprachigen Wörterbuchs für islamische Internatsschüler: Eine korpusgesteuerte Methode

Die Entwicklung dreisprachiger Wörterbücher unter Verwendung von Korpusdaten hat sich als zentrale Methode zur Verbesserung des Sprachenlernens etabliert. Die Erstellung eines dreisprachigen Wörterbuchs für Englisch, Arabisch und Indonesisch, insbesondere im Kontext islamischer Internate, ist jedoch bislang selten vorgekommen. Eine solche Entwicklung wird als notwendig erachtet, um Vorschriften zu berücksichtigen, die es den Schülern verbieten, Wörter online nachzuschlagen. Um ein dreisprachiges Wörterbuch systematisch zu erstellen, das den Sprachgebrauch authentisch widerspiegelt, wurde in dieser Studie eine 4-D-Design-Methodik angewendet: Definieren, Entwerfen, Entwickeln und Verbreiten. Dieser Prozess umfasste die Ermittlung der Bedürfnisse von 50 Schülern und fünf Englischlehrern der Darul Muttaqien NWDI Perian mittels Fragebogen, Interviews und Fokusgruppensitzungen (FGD). Die Ergebnisse zeigten, dass die Schüler sowohl bei allgemeiner als auch bei pesantren-spezifischer englischer Terminologie mit erheblichen Wortschatzlücken konfrontiert sind, was in erster Linie auf begrenzte Sprachkontakt- und Übungsmöglichkeiten zurückzuführen ist. Daher war die Entwicklung eines dreisprachigen Wörterbuchs nach den Ergebnissen der Bedarfsanalyse unerlässlich. Diese Studie ging über den unmittelbaren Nutzen der Wortschatzanreicherung hinaus und schlug zudem einen Paradigmenwechsel hinsichtlich der Art und Weise vor, wie dreisprachige Wörterbücher im Bildungskontext, insbesondere in islamischen Internaten, erstellt und genutzt werden können. Durch die Nutzung von Korpusdaten können

Lehrkräfte dynamischere und anpassungsfähigere Sprachressourcen schaffen, die den sich wandelnden Bedürfnissen der Schüler gerecht werden.

Schlüsselwörter: Dreisprachiges islamisches Wörterbuch, Wortschatzerweiterung, Korpusgesteuerte Methodik, Vier-D-Design

Authors: Yulia Agustina, University Hamzanwadi, Jl. TGKH. M. Zainuddin Abdul Madjid 132 Pancor, East Lombok, West Nusa Tenggara, Indonesia, e-mail: rahestin@gmail.com

M. Adib Nazri, University Hamzanwadi, Jl. TGKH. M. Zainuddin Abdul Madjid 132 Pancor, East Lombok, West Nusa Tenggara, Indonesia, e-mail: adibnazri88@gmail.com

Muhammad Husnu, University Hamzanwadi, Jl. TGKH. M. Zainuddin Abdul Madjid 132 Pancor, East Lombok, West Nusa Tenggara, Indonesia, e-mail: oesnoedotuu@gmail.com

Received: 14.12.2024

Accepted: 15.3.2026

1. Introduction

The expansion of English and Arabic education in Indonesian Islamic boarding schools, known as pesantren, has become increasingly significant. Traditionally focused on religious and moral education, these institutions are now incorporating foreign language instruction into their curricula to prepare students for a globalized world. The integration of English and Arabic not only encourages language proficiency but also promotes cultural engagement with diverse perspectives, which is crucial in today's interconnected society (Ammar et al. 2022: 183). This shift reflects a broader trend in educational reform aimed at equipping students with essential skills for effective communication in both local and international contexts (Huang 2023: 719).

Central to this linguistic competence is a robust vocabulary, which serves as the foundation for effective communication. Despite its importance, students in pesantren often experience restricted linguistic growth due to minimal immersion and lack of exposure outside the formal classroom. Previous research (Miranda/Wahyudin 2023: 41) has consistently shown that vocabulary mastery directly correlates with overall language proficiency, influencing all four language skills: listening, speaking, reading, and writing. The broader an individual's vocabulary, the more adept they become at articulating ideas and understanding others (Khan et al. 2018: 414). Thus, vocabulary acquisition remains a critical focus in language education, especially in specialized environments (Naqsyabandiyah/Dehghanitafti 2023: 38).

However, pesantren students face unique challenges, such as difficulty in memorizing terms, inconsistencies between written and spoken forms, and infrequent use of English leading to vocabulary loss (Agustina et al. 2023: 374). These issues are often exacerbated by reliance on traditional teaching methods and a lack of access to modern learning resources (Zuhairi/Mistar 2023: 454). Furthermore, pedagogical approaches frequently fail to meet the diverse needs of students, hindering long-term retention and application (Sawaki 2017: 1). Addressing these gaps requires the development of

specialized linguistic tools that align with the rigorous principles of practical lexicography (Atkins/Rundell 2008: 45). In a multilingual pesantren environment, a specialized approach is necessary to bridge the gap between English, Arabic, and Indonesian (Hartmann 2001: 58).

The application of corpus linguistics presents a promising avenue for enhancing vocabulary acquisition. By leveraging corpus data, educators can provide learners with authentic language examples that illustrate contextual word usage, facilitating a deeper understanding of vocabulary (Nazeer et al. 2023: 1). This methodology promotes active learning and improves the retention of new terminology (Đurović/Bauk 2022: 335). Additionally, technology-assisted corpus software can support various learning styles through interactive resources.

Yet, a significant paradox exists: while technology offers these benefits, strict regulations in many pesantren regarding digital devices limit students' access. These regulations aim to create an environment conducive to spiritual growth, shielding students from the distractions of modern technology (Ma'ruf et al. 2024: 2473). The pesantren system prioritizes moral and religious discipline, which is often perceived as being compromised by unrestricted digital access (Zukdi et al. 2022: 1142). Consequently, there is an ongoing debate about the need for pesantren to adapt their policies to balance traditional teachings with modern educational tools (Adiyono et al. 2024: 289).

In light of these constraints, the researchers initiated the development of a corpus-based trilingual dictionary in printed form. Unlike electronic resources, manual dictionaries encourage students to engage more deeply with the language through active searching, which aids in memorization and structural understanding (Raihana 2023: 3). While generic trilingual dictionaries are available, they often fail to capture the specialized lexicon and collocations unique to the pesantren system, such as specific religious terminology and social registers. This study addresses this void by creating a context-specific linguistic tool that aligns with the students' lived experiences (Grini 2023: 27).

Utilizing corpus data ensures that the dictionary reflects contemporary language use. Studies in phraseological and terminological lexicography have demonstrated that corpus-based methodologies enhance the accuracy and relevance of dictionary entries by aligning words across different languages through parallel corpora (Gizatova 2018: 357, Liu et al. 2015: 1, Shoba 2018: 2). Despite this potential, the creation of a specialized English-Arabic-Indonesian dictionary for pesantren remains infrequent. This research aims to address this gap by answering three fundamental questions:

- In what ways can a corpus-based methodology contribute to the development of a specialized trilingual dictionary for specific domains?
- How is the proposed trilingual dictionary to be developed?
- What are the implications of corpus data for language learning and translation practices?

2. Method

2.1 Research design

This study employed the 4-D model (Define, Design, Develop, and Disseminate) developed by Thiagarajan, selected for its practical and systematic approach to product development (Santoso et al. 2021: 133, Rais et al. 2022: 11825). The research stages are summarized in Figure 1.

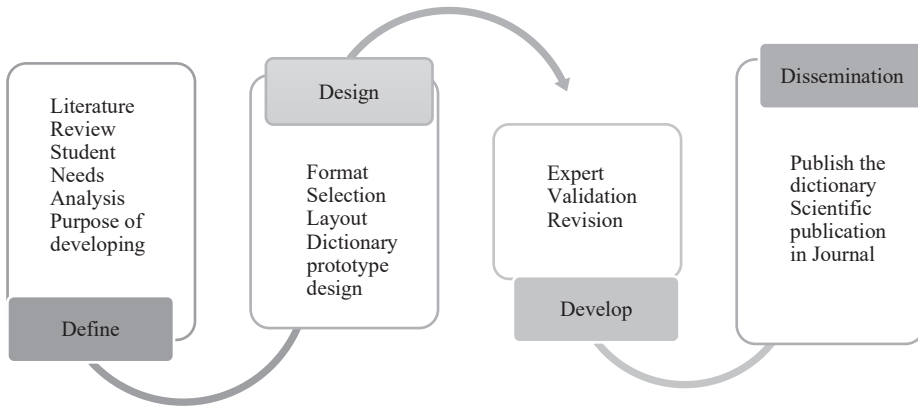


Fig. 1. Research method

First, the Define phase involved reviewing relevant literature related to corpus-based dictionary development and determining the specific needs and target audience for the trilingual dictionary, which was crucial for ensuring that it addressed the students' needs (Luporini 2023: 35). Corpus data analysis helps in understanding the frequency and context of word usage, which guide the selection of entries and definitions (Altun 2021: 511, Gizatova 2018: 357).

To identify specific needs, 50 students and five English teachers from Darul Muttaqien NWDI Perian were selected via purposive sampling, as the school already integrates English and Arabic into daily interactions. Data were collected through a 10-item, four-point Likert-scale questionnaire and Focus Group Discussions (FGD) with teachers (Arifin et al. 2023: 146, Agustina et al. 2023: 375). While questionnaire data were analyzed using descriptive statistics, the FGD results followed the interactive model of Miles, Huberman, and Saldaña (2014: 14). To ensure linguistic and semantic precision, the draft was rigorously validated by a senior linguist and a **Syeikh** (religious authority). The Syeikh's expertise in Arabic philology was crucial for verifying that the trilingual equivalents were not only grammatically correct but also contextually appropriate for the pesantren environment, particularly in correcting subtle semantic nuances.

In the Design phase, the dictionary's structure and layout were conceptualized by integrating trilingual entries with contextual examples (Liu et al. 2015: 1). The researchers adopted Hartmann's (2001: 58) principles, focusing on both macrostructure

(systematic lemma selection) and microstructure (internal entry data, including grammatical labels and corpus-derived translations). This dual-layer design ensures the dictionary functions as a pedagogical resource rather than a mere translation tool (Tarp 2008: 11). Lexical data was extracted from a specialized corpus of 49,948 tokens following procedures established by Agustina et al. (2024: 806). By employing a corpus-based lexicography approach (Granger 2012: 14), the researchers identified high-frequency words to ensure that the trilingual equivalents reflect authentic language usage relevant to the students' daily pesantren lives.

During the Development phase, the trilingual dictionary underwent a formal validation process to ensure it met professional lexicographical standards. The two experts focused on verifying the semantic equivalence across English, Arabic, and Indonesian. Their review was instrumental in refining the draft, ensuring that the final content was both academically rigorous and contextually precise for the target users

Last, the Dissemination phase focused on the distribution of the dictionary, in printing formats and publishing the article, along with strategies for user engagement and feedback collection to continuously improve the dictionary based on user experiences (Tran et al. 2022: 78934). This comprehensive approach not only enhances the quality of the trilingual dictionary but also ensures its relevance and accessibility to the students.

2.2 Corpus design and data collection procedures

To ensure the trilingual dictionary authentically reflects the linguistic needs of Islamic boarding school students, this study employed a corpus-based approach for vocabulary selection. The specialized corpus was systematically compiled from various sources, including English textbooks specifically designed for the Islamic boarding school curriculum, several articles related to the pesantren context, and a collection of students' speech texts. This process resulted in a corpus consisting of approximately 49,957 total words (tokens) with 3,745 unique words (types). While Voyant Tools was utilized for initial computational processing, a manual linguistic analysis was subsequently performed to filter the data based on three primary criteria: high frequency of occurrence, contextual relevance to daily boarding school life, and thematic breadth across different source texts. This systematic extraction procedure ensures that the final lexicographic product is grounded in empirical language data rather than general educational materials.

3. Results and Discussion

3.1 Corpus-based methodologies contributing to the development of specialized trilingual dictionaries

Corpus-based methodologies have significantly transformed the development of specialized trilingual dictionaries by enabling lexicographers to utilize large datasets for

more accurate and contextually relevant entries. The integration of corpus linguistics into dictionaries allows for the extraction of real-world language use, crucial for creating dictionaries that reflect contemporary usage trends. For instance, Gizatova (2018: 357) discusses the construction of an English-Russian phraseological dictionary based on corpus data, emphasizing the importance of analyzing idioms and their frequency to ensure the dictionary's comprehensiveness and representation of actual language use (357). This approach improves trilingual dictionary quality, resolves semantic asymmetry between languages, and enhances usefulness for translators and language learners.

Moreover, Chen's work on the Dictionary and Corpus of Phraseology (DiCoP) highlights the necessity of integrating phraseographic microstructures that account for cultural nuances in language (Chen 2022: 23). This is particularly relevant in specialized fields where terminology can vary significantly across cultures. By employing a corpus-based methodology, lexicographers can ensure that trilingual dictionaries provide translations and cultural context for better term understanding (Chen 2022: 23). This approach aligns with the findings by Shoba (2018: 2) that the use of parallel corpora in developing specialized bilingual dictionaries can enrich the lexicographic process. The effectiveness is further enhanced by the use of electronic and online resources, which facilitate more convenient access to a wider range of dictionary types. Translation processes often rely on various types of dictionaries, including specialized trilingual ones (Altuwairesh 2021: 74). The shift towards electronic types has made navigating complex terminologies easier, as they often include features such as search functions and hyperlinked entries that traditional dictionaries lack (Altuwairesh 2021: 74). This technological advancement supports the corpus-based methodology by allowing lexicographers to continuously update and refine entries based on user interactions and emerging language trends.

Finally, the role of a corpus-based methodology in specialized trilingual dictionary development was underscored by the need for ongoing study and adaptation to changing linguistic landscapes. As noted by Abdelzaher (2022: 163), the contributions of corpus linguistics to lexicographic challenges are increasingly recognized, with many specialized dictionaries now employing corpus tools to enhance their entries. This dynamic process of dictionary creation enhances trilingual dictionaries' accuracy and relevance, ensuring their value for learners and professionals. By embracing a corpus-based methodology, lexicographers can better address the complexities of language and provide users with dictionaries that are both practical and reflective of real-world usage.

3.2 Lexicographical structure of the trilingual dictionary

To address the need for a structured learning tool, the developed trilingual dictionary follows a systematic lexicographical microstructure. Each word entry was designed to provide comprehensive linguistic information, including the English headword, its grammatical category (part of speech), and its equivalent in Arabic and Indonesian. Furthermore, a phonetic transliteration for the Arabic terms is included to assist

students in correct pronunciation. To ensure contextual understanding, each entry was accompanied by an illustrative sentence derived directly from the pesantren corpus. The structural representation and internal information of a dictionary entry, demonstrating the integration of corpus-derived context, are exemplified in Table 1 below.

Component	Description / Example
Headword (English)	Sermon
Part of Speech	Noun (n.)
Arabic Equivalent	كُتْبَاطِح (Khutbah)
Indonesian Equivalent	Khotbah / Ceramah Agama
Corpus Reference	Specialized English Pesantren Corpus, 2024

Tab. 1. Representative example of a trilingual dictionary entry

The practical implementation of this framework in the printed layout of the dictionary was further demonstrated in Figure 2. This figure illustrates the visual typography and arrangement of the headword, word class, and trilingual equivalents as they appear in the book. This design ensures a user-friendly and pedagogically sound reference, specifically tailored for students’ daily academic and religious activities within the boarding school environment.

English	Arabic (العربية)	Indonesian
Advancement (Noun)	التَقَدُّم	Kemajuan
Advances (Noun, Verb)	التَقَدُّمَات	Kemajuan, Memajukan
Advantage (Noun)	المِيزَةُ	Keuntungan
Advantageous (Adjective)	مِيزَةُ	Menguntungkan
Advantages (Noun)	مِيزَاتِنَا	Keuntungan
Advice (Noun)	نَصِيحَة	Saran
Advise (Verb)	نَصَحَ	Menyarankan
Advocate (Noun, Verb)	نَصِيحَةٌ	Pendukung, Menganjurkan
Aesthetic (Adjective, Noun)	الجَمَالِيَّة	Estetik, Estetika

Fig. 2. Page layout and entry presentation of the printed trilingual dictionary

The dual-expert validation provided a comprehensive evaluation, bridging Arabic religious nuances with English linguistic structures. While the Syekh ensured the precision of classical Arabic synonyms, the second validator verified the accuracy of English headwords and grammatical categories. This rigorous process significantly enhanced the dictionary’s reliability, effectively addressing potential inaccuracies from the preliminary drafts.

As shown in Table 2, the average score was 4.1, which falls into the Recommended category. This score reflects that while the dictionary is suitable for use, there were still

areas that required minor adjustments based on the experts’ feedback, particularly regarding the choice of more common synonyms for students.

Expert Validator	Area of Expertise	Score (1-5)	Status / Category
Expert 1	Arabic and Religious Context	4.2	Recommended
Expert 2	English Language and Vocabulary	4.0	Recommended
Average Score		4.1	Recommended with Minor Revisions

Tab. 2. Summary of expert validation results

3.3 A trilingual dictionary development

3.3.1 Define

The results from the literature review, analysis of students’ needs, and the purpose of developing the dictionary have been previously documented (Agustina et al. 2023: 806, Agustina et al. 2024: 376); however, the specific quantitative findings are summarized in Table 2. Table 2 outlines the students’ learning needs and challenges, focusing on English language proficiency within an Islamic boarding school context. The data reveals that key challenges include a lack of effective educational media, difficulties with pronunciation, and limited mastery of English terminology specific to pesantren life. The analysis underscores a significant demand for varied pedagogical tools, such as a specialized dictionary, to maintain engagement and facilitate vocabulary memorization a primary requirement for daily communication.

No	Statements	Likert Scale				% Agreement (Scale 3+4)
		1	2	3	4	
a. Target Needs						
1	The lesson should prevent boring procedures by using varied media.	0	6	14	30	88 %
2	The class should help students become proficient in pesantren-relevant language	5	4	12	27	82 %
3	I would like to learn a large amount of vocabulary.	0	6	16	28	88 %
4	Students want to speak English and memorize specialized vocabulary.	1	5	20	24	88 %
5	My English command is lacking, particularly in pesantren terms.	0	4	21	25	92 %
6	I need media such as dictionaries to improve my vocabulary.	3	3	14	30	88 %
7	I have problems comprehending English texts.	5	3	20	22	84 %
8	Difficulties are mainly in pronunciation and vocabulary knowledge.	0	5	20	25	90 %
	The learning process is not adequately supported by educational media.	4	30	12	4	32 %
b. Learning Needs						
9	Learning should improve general and pesantren-specific vocabulary.	2	28	12	8	40 %
10	To assist students in communication skills and scholarship readiness.	1	23	14	12	52 %

Tab. 2. Students’ needs analysis

Data in Table 2 reveals a critical demand for specialized linguistic support. A significant majority of students (92 %) identified a lack of pesantren-specific English terminology, while 88 % emphasized the necessity of dictionaries for vocabulary acquisition. In contrast, only 32 % felt the current learning process was adequately supported by existing media. These findings establish a strong empirical basis for a corpus-based trilingual dictionary to bridge identified lexical gaps. To complement these student perspectives, Table 3 presents teachers' observations regarding communicative needs and required linguistic resources.

Sources	Thematic Category	Key Statements
T1, T3, T5	Vocabulary Mastery	Students must master extensive vocabulary to communicate effectively in daily and academic life.
T2, T4	Resource Provision	Management must provide attractive media, such as dictionaries, to facilitate vocabulary acquisition.

Tab. 3. Teachers' perspectives on pedagogical needs

Additionally, the results of the teachers' interviews (Table 3) highlight a consensus on the importance of students' communicative drive in both classroom activities and daily interactions. Mastery of English vocabulary is identified as a critical goal, necessitating robust support through resources like a trilingual dictionary. Overall, fostering a wide range of general and pesantren-specific terminology is deemed essential. By providing a structured lexical resource, the gap between the students' current proficiency and the sociolinguistic demands of their environment can be systematically addressed.

3.3.2 Design

The second stage, Design, focused on transforming the needs analysis data into a structured lexicographic draft. In this stage, the lemma selection process was not merely frequency-based but was informed by the "user-research" paradigm in metalexigraphy. Following Tarp's (2008: 11) theory on lexicographic functions, the headwords were selected based on the "communicative needs" of pesantren students in specific situational contexts, ensuring the dictionary addresses real-world lexical gaps identified in the initial needs analysis. This stage also involved determining the macrostructure (the A-Z arrangement) and the microstructure (the internal information of each entry).

To implement these theoretical principles, three primary activities were conducted. First, the results of the needs analysis were analyzed to define the dictionary's scope. Second, vocabulary entries were crafted by drawing from the Islamic boarding school English corpus previously created using the Voyant tool (Agustina et al. 2024: 809). This lexical extraction involved technical stages to ensure the quality of the macrostructure. From the total data, the researchers performed stop-word removal to filter out common functional words (e. g., *the, is, in, at*) that do not carry specific pedagogical value. The extraction then focused on high-frequency keywords and content words specifically related to the pesantren environment.

This filtering process successfully narrowed down the unique word types, which were then manually reviewed to select the final headwords. This manual filtering was guided by a pedagogical utility framework: words with high frequency but low contextual relevance (e. g., technical industrial terms) were excluded. Conversely, medium-frequency words appearing across multiple source texts, such as textbooks and student speeches, were prioritized to ensure thematic breadth. Finally, the Arabic and Indonesian translations were adjusted through collaboration with proficient students to complete the initial prototype.

3.3.3 Develop

The aim of this stage was to modify and develop the trilingual dictionary through validation, revision, and a limited trial. To evaluate the truthfulness of the product, two language professionals with Arabic and English instruction background verified the product. The product was updated following the feedback, with the changes focusing on the meaning, punctuation, and word classes in the Arabic part. This rigorous dual-layer vetting minimized the risk of “translation loss” commonly found in student-developed drafts. The product field testing was conducted in the school setting, serving as the media in learning English.

In the development phase, the data extracted from the corpus was transformed into organized dictionary entries. The lexicographic structure of the dictionary was designed to provide comprehensive linguistic support. Each entry consists of: (1) the English headword derived from the corpus analysis, (2) its part of speech (e. g., noun, verb), (3) the Arabic equivalent, (4) the Indonesian translation, and (5) a contextual example sentence to demonstrate practical usage. This structured format ensures that students not only understand the meaning of a word but also know how to apply it in their daily communication within the pesantren environment.

3.3.4 Disseminate

The results of this study were disseminated in local and international contexts. A local dissemination has already been conducted in Darul Muttaqien NWDI Perian while some printed dictionaries will be distributed to other boarding schools in the year of 2025. Another way of dissemination is by publishing the results in an international journal.

3.4 The implications of corpus-based methodologies for language learning and translation practices

The implications of corpus data for language learning and translation are profound. Corpus linguistics provides a robust empirical foundation for understanding contextual language use, allowing learners to grasp authentic patterns, collocations, and idiomatic expressions (Demyanchuk 2023: 18, Chen 2022: 23). This exposure enhances lexical awareness and a nuanced understanding of linguistic functions across various

communicative situations. While communicative competence encompasses broader sociocultural dimensions, this corpus-based dictionary serves as a crucial mediating tool in the restricted pesantren environment. It provides the essential “building blocks” vocabulary and collocations required for students to eventually achieve full proficiency.

In translation practice, corpus tools enhance efficiency by providing a rich lexical context for informed choices (Meng 2021: 33, Sonbul et al. 2024: 17). This study specifically addresses semantic asymmetry, such as mapping English abstract nouns to Arabic religious equivalents. For instance, corpus-derived illustrative sentences for Sermon (*Khutbah*) demonstrate how terms function differently in religious versus general academic settings. Furthermore, this integration fosters autonomous learning and critical thinking, empowering students to analyze complex terminologies independently (Vojnović 2021: 53, Solopova/Khomutova 2022: 1056). Ultimately, these innovative pedagogical approaches prepare students for the challenges of real-world, context-sensitive translation.

4. Conclusion

The integration of corpus-based methodologies into trilingual dictionary development significantly enhances relevance and utility for users by leveraging real-world language usage. This approach enables lexicographers to create entries reflecting contemporary linguistic practices, ensuring learners receive pertinent vocabulary and contextual data. Such methodology enriches content, meeting the needs of learners and professionals who require precise terminology and a deeper understanding of collocations and idiomatic expressions.

While this study is confined to an English-Arabic-Indonesian pair within an Islamic boarding school context, its implications are far-reaching. Beyond its immediate use, this corpus-based framework serves as a scalable model for developing specialized dictionaries in other vocational or restricted-technology environments. Future research should investigate a wider variety of languages and long-term impacts on language acquisition. Furthermore, future iterations could explore transitioning from printed formats to offline digital applications, ensuring pedagogical innovation respects the disciplinary boundaries of the pesantren system. Overall, this analysis underscores the importance of integrating corpus data into lexicography, paving the way for more effective, user-centered trilingual dictionaries that meet the evolving demands of language learners and translators.

Literature

ABDELZAHER, Esra Moustafa. „An Investigation of Corpus Contributions to Lexicographic Challenges Over the Past Ten Years“. *Lexikos* 32 (2022): 162–79. <https://doi.org/10.5788/32-1-1714>. 25.11.2024.

- ADIYONO, Syamsun Ni'am and Ahmad ANSHOR. „Islamic Character Education in the Era of Industry 5.0: Navigating Challenges and Embracing Opportunities“. *AL-HAYAT Journal of Islamic Education* 8 (2024): 287–304. <https://doi.org/10.35723/ajie.v8i1.493>. 1.9.2024.
- AGUSTINA, Yulia, Widodo PRATOMO and Margana MARGANA. „Creating an Islamic Boarding School English Corpus: Corpus Metadata, Frequently Used Words, and Unique Words“. *Journal of Education and Learning (EduLearn)* 18 (2024): 804–816. <https://doi.org/10.11591/edulearn.v18i3.21095>. 25.11.2024.
- AGUSTINA, Yulia, Margana MARGANA, Putro NUR and Adib NAZRI. „Designing English Instructions for Islamic Settings: A Need Analysis in Indonesian Pesantren“. *World Journal of English Language* 13 (2023): 373–382. <https://doi.org/10.5430/wjel.v13n8p373>. 20.6.2024.
- ALTUN, Hatice. „The Learning Effect of Corpora on Strong and Weak Collocations: Implications for Corpus-Based Assessment of Collocation Competence“. *International Journal of Assessment Tools in Education* 8 (2021): 509–526. <https://doi.org/10.21449/ijate.845051>. 5.11.2024.
- ALTUWAIRESH, Nasrin. „Translation Trainees' Uses of and Attitudes Towards Different Types of Dictionaries: A Case Study“. *International Journal of Applied Linguistics & English Literature* 10 (2021): 73–80. <https://doi.org/10.7575/aiac.ijalel.v10n.2p.73>. 25.11.2024.
- AMMAR, Adzfar, Abdul MUNIP, Musthofa TULUS, Ahmad SAPAR and Agung SETIYAWAN. „The Contribution of Pesantren's Meaning Symbols and Their Effects on Translation of Arabic Text“. *Arabi : Journal of Arabic Studies* 7 (2022): 182–197. <https://doi.org/10.24865/ajas.v7i2.450>. 6.1.2025.
- ANIS, Grini. *The Impact of Using Bilingual Dictionaries on Literary Translation*. The University of Guelma, 2023. <https://dspace.univ-guelma.dz/jspui/bitstream/123456789/15566/1/M821.421.pdf>. 10.9.2024.
- ARIFIN, Yulyani, Elizabeth GUNAWAN and Margaretha OHYVER. „Developing an Online Information Portal for Enhancing Society Awareness of RPTRA (a Case Study RPTRA MAYA ASRI 13)“. *Procedia Computer Science* 216 (2023): 144–150. <https://doi.org/10.1016/j.procs.2022.12.121>. 2.1.2025.
- ATKINS, B. T. Sue and Michael RUNDELL. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press, 2008.
- CHAN, Alice Yin Wa. „The Effectiveness of Using a Bilingualized Dictionary for Determining Noun Countability and Article Selection“. *Lexikos* 27 (2017): 183–213.
- CHEN, Lian. „Phraseoculture in the Construction of the Corpus of the DiCoP: The Treatment of the Phraseographic Microstructure“. *EUROPHRAS: International Conference Computational and Corpus-based Phraseology* (2022): 17–25.
- DEMYANCHUK, Yuliya. „Term-Combination in the Context of Corpus-Applied Translation Studies“. *Polonia University Scientific Journal* 59 (2023): 16–22.
- ĐUROVIĆ, Zorica and Sanja BAUK. „Testing Corpus Linguistics Methods in ESP Vocabulary Teaching“. *Journal of Teaching English for Specific and Academic Purposes* 10 (2022): 319–337.
- GIZATOVA, Guzel. „A Corpus-Based Approach to Lexicography: A New English-Russian Phraseological Dictionary“. *International Journal of English Linguistics* 8 (2018): 357.
- GRANGER, Sylviane. „Electronic Lexicography: From Digitization to Corporatization“. *The Oxford Handbook of Lexicography*. Ed. Philip Durkin. Oxford: Oxford University Press, 2012.
- GRINI, Mohammed Anis. *The Impact of Using Bilingual Dictionaries on Literary Translation. Disertation* (2023). <https://dspace.univ-guelma.dz/jspui/bitstream/123456789/15566/1/M821.421.pdf>
- HARTMANN, Reinhard. R. K. *Teaching and Researching Lexicography*. London: Routledge, 2001. Print.

- HUANG, Shuping. „Incorporating Data-Driven Learning Approach Into English Language Teaching by Using Corpus-Based Tools“. *Advances in Intelligent Systems Research* (2023): 713–720.
- KHAN, Raja, Radzuan NOOR, Muhammad SHAHBAB, Ibrahim AINOL and Mustafa GHULAM. „The Role of Vocabulary Knowledge in Speaking Development of Saudi EFL Learners“. *Arab World English Journal* 9 (2018): 406–418.
- LIU, Xiaodong, Kevin DUH and Yuji MATSUMOTO. „Multilingual Topic Models for Trilingual Dictionary Extraction“. *ACM Transactions on Asian and Low-Resource Language Information Processing* 14 (2015): 1–22.
- LUPORINI, Antonella. „Exploring the Lexis of Art Through a Specialized Corpus: A Bilingual Italian-English Perspective“. *International Journal of English Linguistics* 13 (2023): 29–38. <https://doi.org/10.5539/ijel.v13n7p29>. 10.11.2024.
- MA, Qing and Fang MEI. „Review of Corpus Tools for Vocabulary Teaching and Learning“. *Journal of China Computer-Assisted Language Learning* 1 (2021): 177–190.
- MA'RUF, Rachmad Arif, Darmanto DARMANO, Mohamad HADITIA, Fatimatul ASRORIAH, Bima ASY'ARIE and Zuhairi ZUHAIRI. „Islamic Boarding Schools and Technology: Efforts to Overcome Social Changes in Santri Misuse of Gadgets“. *AL-ISHLAH Jurnal Pendidikan* 16 (2024): 2472–2484.
- MENG, Qingliang. „The Pedagogy of Corpus-aided English-Chinese Translation From a Critical & Creative Perspective“. *Theory and Practice in Language Studies* 11 (2021): 29–42.
- MILES, Matthew, Michael HUBERMAN and Johnny SALDAÑA. *Qualitative Data Analysis: A Methods Sourcebook*. Los Angeles, London, New Delhi, Singapore, Washington: Sage Publication, Inc., 2014.
- MIRANDA, Junnisa and Achmad WAHYUDIN. „Pre-Service Teachers Strategies in Improving Students' Speaking Skills“. *Journal of English Language Teaching and Learning (JELTL)* 4 (2023): 40–47.
- NAQSYABANDIYAH, Noviana and Najmeh DEHGHANITAFI. „Developing Task-Based Learning Materials to Improve Students' Vocabulary Mastery Viewed From Linguistic Awareness“. *Journal of Language and Literature Studies* 3 (2023): 37–52.
- NAZEER, Imran, Shahira MUKHTAR and Behjat AZHAR. „Exploring the Effectiveness of Vocabulary Acquisition Strategies in Foreign Language Learning“. *Harf-o-Sukhan* 7 (2023): 1–14.
- NINGSIH, Yuliana, Adhiela SYAIEF, Kurnia ARTIKA and Herpendi HERPENDI. „Developing Multilingual Automotive E-Dictionary Based on Corpus Linguistics“. *Advances in Social Science, Education and Humanities Research/Advances in Social Science, Education and Humanities Research* 647 (2022): 962–967.
- RAIHANA, Rana. *The Influence of Language Regulation toward English Learning in Boarding School*. Thesis, UIN Ar-Rainiry, 2023.
- RAIS, Syafruddin, Anni FARIDAH, Yuliana YULIANA and Ambiyar AMBIYAR. „Model Development of Community-based and Digital Tourism Village in Belakang Padang“. *International Journal of Health Sciences* 6 (2022): 11822–11835.
- SAHIN, Abdullah. „Critical Issues in Islamic Education Studies: Rethinking Islamic and Western Liberal Secular Values of Education“. *Religions* 9 (2018): 335.
- SANTOSO, Gunawan Sulisty Budi and H. SUMARJO. „Pengembangan Modul Teknik Pengukuran Tanah untuk Siswa Kelas X Program Keahlian Desain Permodelan dan Informasi Bangunan di SMK Negeri 1 Purworejo“ [Development of the Land Measurement Techniques Module for 10th Grade Students of the Building Design and Information Modeling Expertise Program at SMK Negeri 1 Purworejo]. *Jurnal Pendidikan Teknik Sipil* 3 (2021): 131–143.
- SAWAKI, Yasuyo. „University Faculty Members' Perspectives on English Language Demands in Content Courses and a Reform of University Entrance Examinations in Japan: A Needs Analysis“. *Language Testing in Asia* 7 (2017): 1–16.

- SHOBA, Feziwe. *Exploring the Use of Parallel Corpora in the Compilation of Specialized Bilingual Dictionaries of Technical Terms: A Case Study of English and IsiXhosa*. University of South Africa, 2018.
- SOLOPOVA, Olga Aleksandrovna and Tamara Nikolaevna KHOMUTOVA. „An Explanatory Combinatorial Dictionary of English Conflict Lexis: A Case Study of Modern Political Discourse“. *Russian Journal of Linguistics* 26 (2022): 1050–1077. <https://doi.org/10.22363/2687-0088-32005>. 10.11.2024.
- SONBUL, Suhad, Dina EL-DAKHS and Rezan ALHARBI. „Rendering Natural Collocations in a Translation Task: The Effect of Direction, Congruency, Semantic Transparency, and Proficiency“. *International Journal of Applied Linguistics* 34 (2024): 117–133.
- TARB, Sven. *Lexicography in the Borderland between Knowledge and Communication: Foreign-Language Lexicography and the General Theory of Lexicography*. Berlin, New York: Walter de Gruyter, 2008. Print.
- TRAN, Phuoc, Thien NGUYEN, Dinh-Hong VU, Huu-Anh TRAN and Bay VO. „A Method of Chinese-Vietnamese Bilingual Corpus Construction for Machine Translation“. *IEEE Access* 10 (2022): 78928–78938.
- UCHIHARA, Takumi and Jon CLENTON. „Investigating the Role of Vocabulary Size in Second Language Speaking Ability“. *Language Teaching Research* 24 (2020): 540–556.
- VOJNOVIĆ, Dragana Vuković. „Key Noun + Noun Collocations in the Language of Tourism: A Corpus-Based Study of English and Serbian“. *ELOPE English Language Overseas Perspectives and Enquiries* 18 (2021): 51–68.
- WIJESURIYA, Keshani and Sucheru DISSANAYAKE. „The Impact of Online Monolingual and Bilingual Dictionaries on Vocabulary Learning“. *SLIIT Journal of Humanities and Sciences* 1 (2020): 62–69.
- YAO, Gang. „Vocabulary Learning Through Data-driven Learning in the Context of Spanish as a Foreign Language“. *Research in Corpus Linguistics* 7 (2019): 18–46.
- ZUHAIRI, Alfian and Junaidi MISTAR. „Vocabulary Learning Strategies and Vocabulary Mastery by Indonesian EFL Learners“. *World Journal of English Language* 13 (2023): 453–461.
- ZUKDI, Iipi, Zulvia TRINOVA, Zulkifli ZULKIFLI, Iman NASURION and Salmiwati SALMIWATI. „The Role of Islamic Boarding Schools In Building The Character of The Nation“. *Edukasi Islam: Jurnal Pendidikan Islam* 11 (2022): 1137–1150.

Author's Declaration on the Use of Generative Artificial Intelligence (GenAI) Tools

During the preparation of this work the authors used [Scite AI] in order to search for sources. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

ZITIERNACHWEIS:

AGUSTINA, Yulia, NAZRI, M. Adib, HUSNU, Muhammad. „Entwicklung eines dreisprachigen Wörterbuchs für islamische Internatsschüler: Eine korpusgesteuerte Methode“, *Linguistische Treffen in Wrocław* 29, 2026 (I): 31–44. DOI: 10.23817/lingtreff.29-2.